The position of water molecules in Bacteriorhodopsin: A three-dimensional distribution function study

Yokogawa, Daisuke; Sato, Hirofumi; Sakaki, Shigeyoshi

The Position of Water Molecules in Bacteriorhodopsin: A Fragment Three-dimensional Reference Interaction Site Model Study

Daisuke Yokogawa, Hirofumi Sato *, Shigeyoshi Sakaki

Department of Molecular Engineering, Graduate School of Engineering, Kyoto University, Nishikyo-ku, Kyoto 615-8510, Japan

Abstract

Water molecules often play a crucial role in biological system and their positions are of primary concern. Our recently developed fragment three-dimensional reference interaction site model (fragment 3D-RISM), which is a highly parallelizable method based on the 3D-RISM theory, was applied to obtain the distribution of water molecules in Bacteriorhodopsin. The bound water molecules in the vicinity of Shiff base were correctly reproduced at a reasonable computational time.

1 Introduction

The information of hydration structure is very fundamental in biosystem [1]. Hydrogen bondings between waters and protein affect protein structure and activity of enzyme. The information is also very useful in drag design since the position of waters in the vicinity of the activity site has great influence on the stabilization of the drug-protein interaction [2; 3] .

The waters in biosystem can be classified into "surface" or "bound", according to whether they are surrounded by other water molecules or protein [**? ?**] For example, it is well known that hydrogen bonded water molecules play a key role in Bacteriorhodopsin (bR), which is a light-driven proton pump in *Halobacterium salinarum*. To reveal the mechanism of the pump, a

* Corresponding author. FAX: +81-75-383-2799
Email address: hirofumi@moleng.kyoto-u.ac.jp (Hirofumi Sato).

huge number of approaches including X-ray crystallography [5], resonance Raman [6] and Fourier transform infrared spectroscopy [7] have been performed. These experimental approaches have elucidated that the hydrogen-bonding networks of these water molecules providing the proton pathway in the pump [7]. Theoretical approaches to study the mechanism have been also performed [8; 9; 10; 11; 12; 13; 14]. In most of the studies, the mechanism of bR function was focused and the initial positions of bound waters inside protein were taken from experimental data.

The theoretical prediction of the water distribution is still very limited. It is extremely difficult task for the molecular simulation such as molecular dynamics to compute the distribution because it is necessary to sample the interaction between the protein and water molecules on the extensive free energy hypersurface. It has been shown that an integral equation theory for molecular liquids, three-dimensional reference interaction site model (3D-RISM) [15; 16], is a powerful tool to study the distribution of bound and surface waters and numerous applications have been carried out [17; 18; 19; 20; 21]. We have recently developed a new approach based on 3D-RISM, in which the equations are elaborated so that the high parallel performance can be achieved (fragment 3D-RISM) [22]. Similar to the original 3D-RISM, our method evaluates the 3D solvation structure but the required computational time can be compressed. Although the present theory is regarded as an approximation of 3D-RISM, the obtained distribution function is virtually the same as that by the original method [22]. Furthermore, the new theory is practically free from the grid size since the distribution functions are computed by the expansion around the individual solute site. Actually the functions are described in higher-resolution than the original one.

In the present work, we applied the fragment 3D-RISM to computations of the waters' position in bR and compared with those obtained by X-ray crystallography as well as by previous simulations. After brief description of the method, details of the calculation are explained in section III. The bound waters are discussed in section IV.

2 Method

3D correlation functions of a solvent site α, H_α and C_α are expressed with reference and residual correlation functions, as follows:

$$H_\alpha(\mathbf{r}) = H_\alpha^{\mathrm{ref}}(\mathbf{r}) + \Delta H_\alpha(\mathbf{r}), \tag{1}$$
$$C_\alpha(\mathbf{r}) = C_\alpha^{\mathrm{ref}}(\mathbf{r}) + \Delta C_\alpha(\mathbf{r}). \tag{2}$$

For the reference correlation functions, H_α^{ref} and C_α^{ref}, 1D correlation functions are employed [22]. The residual functions, ΔH and ΔC, are 3D functions and divided into the components localized on each atomic site of solute,

$$\Delta H_\alpha(\mathbf{r}) = \sum_\beta w_\beta(\mathbf{r}) \Delta H_\alpha^{(\beta)}(\mathbf{r}_\beta), \tag{3}$$

$$\Delta C_\alpha(\mathbf{r}) = \sum_\beta w_\beta(\mathbf{r}) \Delta C_\alpha^{(\beta)}(\mathbf{r}_\beta), \tag{4}$$

where $w_\beta(\mathbf{r})$ is the weight function for solute atomic site β at the position $\mathbf{r}$. In this work, $\Delta H_\alpha^{(\beta)}$ and $\Delta C_\alpha^{(\beta)}$ are expanded with real solid harmonics S_{lm}, as follows:

$$\Delta H_\alpha^{(\beta)}(\mathbf{r}_\beta) = \sum_{lm} \Delta H_{lm,\alpha}^{(\beta)}(r_\beta) S_{lm}(\hat{\mathbf{r}}_\beta), \tag{5}$$

$$\Delta C_\alpha^{(\beta)}(\mathbf{r}_\beta) = \sum_{lm} \Delta C_{lm,\alpha}^{(\beta)}(r_\beta) S_{lm}(\hat{\mathbf{r}}_\beta), \tag{6}$$

where $\hat{\mathbf{r}}_\beta$ is a unit vector with its origin at atom β. The component of residual total correlation function can be approximated by the following equation in analogy with the original RISM theory,

$$\Delta H_{lm,\alpha}^{(\beta)}(r_\beta) \simeq \sum_\delta \Delta C_{lm,\delta}^{(\beta)} * \left[\omega_{\delta\alpha}^V + \rho h_{\delta\alpha}^V\right](r_\beta). \tag{7}$$

To solve Eqs. (5), (6) and (7), we elaborated the following Kovalenko-Hirata (KH) type closure [23],

$$\Delta H_\alpha^{(\gamma)}(\mathbf{r}_\gamma) = \begin{cases} \exp(\chi_\alpha^{(\gamma)}(\mathbf{r}_\gamma)) - 1 - H_\alpha^{\mathrm{ref}}(\mathbf{r}_\gamma) & \text{for } \chi_\alpha^{(\gamma)}(\mathbf{r}_\gamma) \le 0 \\ \chi_\alpha^{(\gamma)}(\mathbf{r}_\gamma) - H_\alpha^{\mathrm{ref}}(\mathbf{r}_\gamma) & \text{for } \chi_\alpha^{(\gamma)}(\mathbf{r}_\gamma) > 0 \end{cases}$$

$$\chi_\alpha^{(\gamma)}(\mathbf{r}_\gamma) = -u_\alpha(\mathbf{r}_\gamma)/k_BT + \left\{H_\alpha^{\mathrm{ref}}(\mathbf{r}_\gamma) + \Delta H_\alpha^{(\gamma)}(\mathbf{r}_\gamma)\right\} - \left\{C_\alpha^{\mathrm{ref}}(\mathbf{r}_\gamma) + \Delta C_\alpha^{(\gamma)}(\mathbf{r}_\gamma)\right\}, \tag{8}$$

where k_B is Boltzmann's factor. $u_\alpha(\mathbf{r}_\gamma)$ is the intermolecular potential function between solute and the solvent site α, which is evaluated on the grid points around solute site γ. Whilst the hypernetted chain (HNC) approximation is a theoretically well-established closure, using KH-type closure was essential in this study because of its numerical stability. It is known that KH closure sometimes greatly underestimates the site-site correlation functions, but we believe that the discussion in the present work is not considerably affected by the choice of closure equations.

The flow chart of the present method is shown in Scheme 1.

3 Computational details

The reference correlation functions were calculated by 1D RISM/KH procedure [24; 25]. Using the converged 1D direct correlation function, the reference correlation functions H_α^{ref} and C_α^{ref} were evaluated. The residual correlation functions were then calculated by Eqs. (5), (6) and (7), coupled with Eq. (8). In this approach, the calculation was performed on logarithmic grid for radial part and the Lebedev grid [26] for angular part. With the grid set, the convolution integral in Eq. (7) can be calculated by spherical Bessel transformation [27]. To reduce the computational cost per one CPU, calculation was parallelized with MPICH2 [28]. The solvation structure was evaluated with Eqs. (1) and (3) from the obtained residual correlation functions. For visualization of 3D solvation structure, VMD software [29] was used.

4 Results and Discussion

The geometry of bR was taken from the PDB data (PDBID: 1c3w) [5] and OPLS parameters [30] were employed. All the water molecules inside the protein in PDB data were removed before the computation. SPC-like model of water was employed [31] with a correction concerning the Lennard-Jones parameters of the hydrogen sites (σ=1.0Å, ϵ=0.056 kcal mol^{-1}). bR is a large molecule and the number of the solute sites is 2221 for the parameter set. However, the required memory size of the present method was about only 850 MB per one CPU, meaning the allowance to perform the computation even with PC cluster. It is noted bR is in the cell membranes in reality, but they are ignored in the the present computation. We believe that the surrounding water or membranes have no effect on the water molecules *inside* the protein.

bR contains all-*trans* retinal, which binds covalently to Lys216 through a protonated Schiff base linkage. The distribution of waters around retinal calculated by the present method and the water molecules obtained by X-ray diffraction data [5] are shown in Fig. 1. Except for Lys216 and retinal, bR is represented by ribbon for the sake of viewability. The scattered green (white) areas shown in the figure indicate where the distribution function of water oxygen (hydrogen) site is greater than the threshold value, 3.2. The bound waters determined by X-ray crystallography are represented by red spheres. Several positions of the comparatively-large area coincide with the experimental data, suggesting that the water distribution is correctly reproduced by the

present method. The distribution of waters are not continuous and intermitted by residues. In the case of aquaporins, the bound waters are continuously distributed through out the channel [32]. This may be the large difference between a pump and a channel.

Let us look at closely the regions **A** and **B**, which are enclosed with dashed lines. Fig. 2 focuses the water distribution in region **A**, around Schiff base. The scattered green areas shown in the figure indicate where the distribution function of water oxygen site is greater than the threshold value, 3.2. To display the strength of hydrogen bondings clearly, distributions of water hydrogen greater than 2.2, 3.2, and 4.2 are respectively shown in Figs. 2(a), 2(b), and 2(c). The positions of bound waters obtained by X-ray diffraction are also shown with dashed line, **Wa**, **Wb** and **Wc** [5]. Conspicuous localized distributions of oxygen and hydrogen are found surrounded by LYS216, ASP85 and ASP212, which coincides with the result obtained by X-ray crystallography. Shibata et al. proposed that waters **Wa** and **Wb** strongly bind with oxygen site of ASP85, and water **Wc** strongly binds with ASP212 from the FTIR studies [33; 34]. These strong hydrogen bondings are found in Fig. 2(c). The broad distributions in Figs. 2(a) and 2(b) indicate the fluctuation of waters because the present calculation was performed at the condition of room temperature. The schematic picture of the bound waters drawn from these figures is illustrated in Fig. 2(d). The thick and thin dotted lines show the strong and weak hydrogen bondings, respectively, which is in good agreement with the network reported by Shibata et al. In the neighborhood of ARG82, the distribution of oxygen site is found (**Wd**) but no hydrogen site can be seen, at least, with the threshold, 4.2. This means the water **Wd** is captured by the residue but its orientation is relatively free compared to aforementioned water molecules, **Wa**, **Wb**, and **Wc**.

Fig. 3(a) shows the distributions of water oxygen (green) and water hydrogen (white) sites in region **B**, upside the retinal. There are two main diffuse solvation structures (distributions **I** and **II**). By X-ray crystallography, **We** and **Wf** are detected in **I** and **II**, respectively. **We** in the distribution **I** links ALA215 (in helix G) and TRP182. Schulten et al. reported another water molecule in the vicinity of **We** (see Fig. 1 in their work [11]), although no water molecule is reported in the X-ray crystallography except for **We** [5]. Fig. 3(a) shows that there are interaction between water and THR178 (thick dotted line), which may correspond to the water reported by Schulten et al. [11]. Another main distribution (**II**) is continued from LYS216 to ASP96 and THR46 in the vicinity of the helix G backbone, which is distorted from standard α-helical conformation [5]. The water **Wf** binds with the carbonyl group of LYS216. Humphrey et al. proposed other two bound waters in this area [10]. The broad green and white distributions correspond to these waters. The hydrogen bonding network of these waters may be understood as depicted in Fig. 3(b).

Both of **I** and **II** make hydrogen bonding network from the Retinal to ASP96 and to THR46, which is consistent with the previous works' conclusions [5; 10; 11; 14].

5 Concluding Remarks

Fragment 3D-RISM was applied to the calculation of the distribution of bound waters in Bacteriorhodopsin. The computed distributions show good agreement with those by X-ray diffraction experiment. The method is highly parallelizable and can sufficiently reduce the required computational cost and time while adequate distribution of water molecules are obtained.

In the neighborhood of the Schiff base, several water molecules captured by residues were found. ASP85 and ASP212 obviously accept hydrogen bonding from neighbor waters because the direction of oxygen-hydrogen bond in water can be discriminated from the distribution. On the other hand, the water near ARG82 is also captured by the residue but rather freely oriented. Near the G-helix backbone to ASP96, two largely continuous distributions were seen. They are consistent with the bound water molecules reported in molecular simulation study.

The fragment 3D-RISM method is highly efficient with capability to predict the solvation structure concerning sufficiently large bio-molecules.

6 Acknowledgments

This work has been financially supported in part by the Grant-in Aid for Scientific Research on Priority Areas "Water and biomolecules" (430-18031019) and by "Molecular Theory for Real Systems" (461). D. Y. thanks the Grant-in Aid for JSPS Fellows. All of the calculations were carried out PRIMEQUEST in the Institute for Molecular Science (IMS), Okazaki Japan. All of them were supported by the Ministry of Education, Culture, Sports, Science and Technology (MEXT) Japan.

References

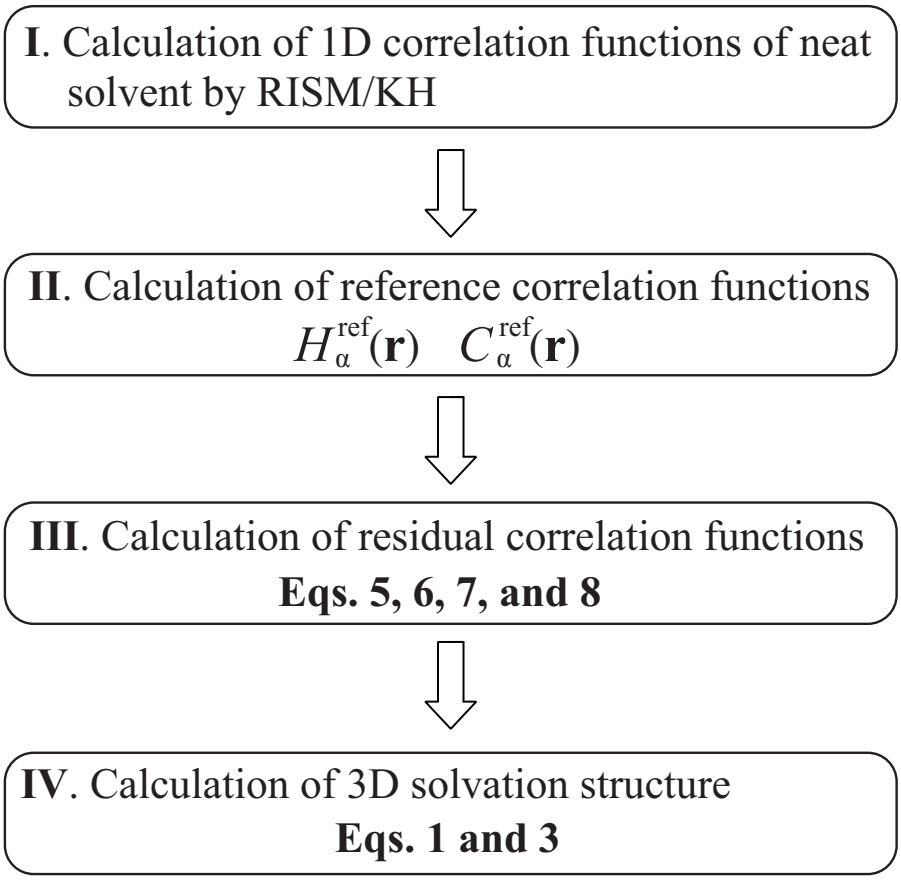
[1] M. Chaplin, Nature Rev. Mol. Cell Biol. 7 (2006) 861.
[2] R. A. Engh, H. Brandstetter, G. Sucher, A. Eichinger, U. Baumann, W. Bode, R. Huber, T. Poll, R. Rudolph, W. von der Saal, Structure 4 (1996) 1353.

[3] J. B. Finley, V. R. Atigadda, F. Duarte, J. J. Zhao, W. J. Brouillette, G. M. Air, M. Luo, J. Mol. Biol. 293 (1999) 1107.
[4] E. Mayer, Protein Sci. 1 (1992) 1543.
[5] H. Luecke, B. Schobert, H.-T. Richter, J.-P. Cartailler, J. K. Lanyi, J. Mol. Biol. 291 (1999) 899.
[6] S. O. Smith, J. Lugtenburg, R. A. Mathies, J. Membr. Biol. 85 (1985) 95.
[7] H. Kandori, Biochim. Biophys. Acta 1460 (2000) 177.
[8] S. Hayashi, I. Ohmine, J. Phys. Chem. B 104 (2000) 10678.
[9] S. Hayashi, E. Tajkhorshid, H. Kandori, K. Schulten, J. Am. Chem. Soc. 126 (2004) 10516.
[10] W. Humphrey, I. Logunov, K. Schulten, M. Sheves, Biochemistry 33 (1994) 3668.
[11] J. Baudry, E. Tajkhorshid, F. Molnar, J. Phillips, K. Schulten, J. Phys. Chem. B 105 (2001) 905.
[12] F. Zhou, A. Windemuth, K. Shulten, Biochemistry 32 (1993) 2291.
[13] M. Nina, B. Roux, J. C. Smith, Biophys. J. 68 (1995) 25.
[14] B. Roux, M. Nina, R. Pomès, J. C. Smith, Biophys. J. 71 (1996) 670.
[15] A. Kovalenko, F. Hirata, Chem. Phys. Lett. 290 (1998) 237.
[16] D. Beglov, B. Roux, J. Phys. Chem. B 101 (1997) 7821.
[17] S. Phongphanphanee, N. Yoshida, F. Hirata, Chem. Phys. Lett. 449 (2007) 196.
[18] T. Imai, R. Hiraoka, T. Seto, A. Kovalenko, F. Hirata, J. Phys. Chem. B 111 (2007) 11585.
[19] T. Imai, R. Hiraoka, A. Kovalenko, F. Hirata, Proteins-Structure Function and Bioinformatics 66 (2007) 804.
[20] T. Imai, R. Hiraoka, A. Kovalenko, F. Hirata, J. Am. Chem. Soc. 127 (2005) 15334.
[21] N. Yoshida, S. Phongphanphanee, F. Hirata, J. Phys. Chem. B 111 (2007) 4588.
[22] D. Yokogawa, H. Sato, T. Imai, S. Sakaki, *to be submitted.*
[23] A. Kovalenko, F. Hirata, J. Chem. Phys. 110 (1999) 10095.
[24] F. Hirata, P. J. Rossky, Chem. Phys. Lett. 83 (1981) 329.
[25] F. Hirata (Ed.), Molecular Theory of Solvation, Kluwer, Dordrecht, 2003.
[26] V. I. Lebedev, D.N. Laikov, Dokl. Math. 59 (1999) 477.
[27] J. D. Talman, J. Comput. Phys. 29 (1978) 35.
[28] MPICH2 High-performance and widely portable implementation of MPI http://www.mcs.anl.gov/research/projects/mpich2/index.php.
[29] W. Humphrey, A. Dalke, K. Schulten, J. Mol. Graphics 14 (1996) 33.
[30] W. L. Jorgensen, OPLS and OPLS-AA Parameters for Organic Molecules, Ions, and Nucleic Acids, Yale University, 1997.
[31] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren and H.J. Hermans, Intermolecular Forces, Ed. B. Pullman (Reidel, Dordrecht, 1981) p. 331.
[32] S. Phongphanphanee, N. Yoshida, F. Hirata, J. Am. Chem. Soc. 130 (2008) 1540.

[33] M. Shibata, T. Tanimoto, H. Kandori, J. Am. Chem. Soc. 125 (2003) 13312.
[34] M. Shibata, H. Kandori, Biochemistry 44 (2005) 7406.

Figure captions

- Fig. 1: 3D distribution of waters inside bR. The green (white) regions correspond to the area where the distribution function of water oxygen (hydrogen) site is larger than 3.2. The bound waters detemined by X-ray crystallography are represented by red spheres.
- Fig. 2: 3D solvation structure of water oxygen (green) and hydrogen (white) in **A**. The positions of bound waters determined by X-ray crystallography are shown with dashed line. The white surface show the areas where the distribution of water hydrogen is larger than (a) 2.2, (b) 3.2 and (c) 4.2. In these panels, distribution of water oxygen larger than 3.2 is shown. In panel (c), possible hydrogen bondings are depicted with dashed line. Schematic drawing of this area is shown in (d).
- Fig. 3: (a) 3D solvation structure of water oxygen (green) and hydrogen (white) in **B**. Distributions of water oxygen and water hydrogen larger than 2.5 are shown. The positions of covalent-bonded hydrogen atoms are explicitly depicted for THR 46 and 178. (b) Schematic drawing of bound waters proposed from panel (a).



Scheme 1

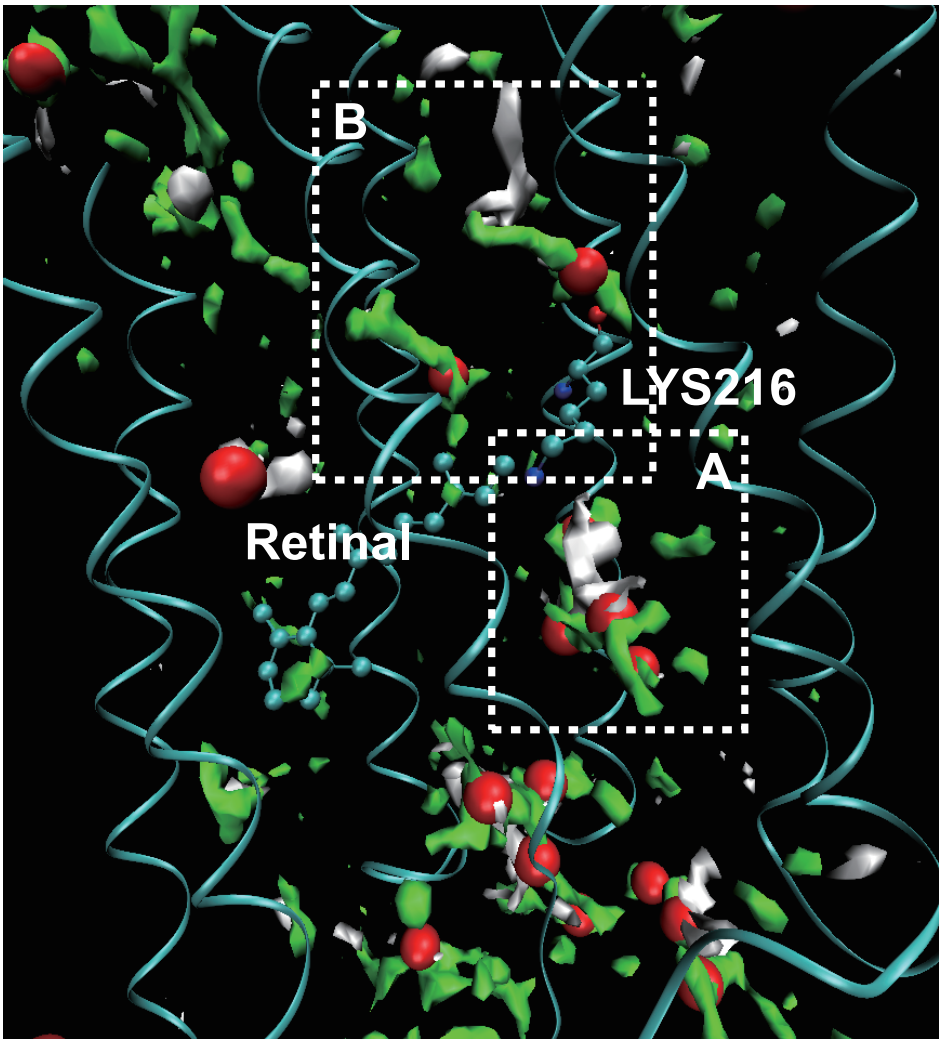


Fig. 1. Yokogawa *et al.*

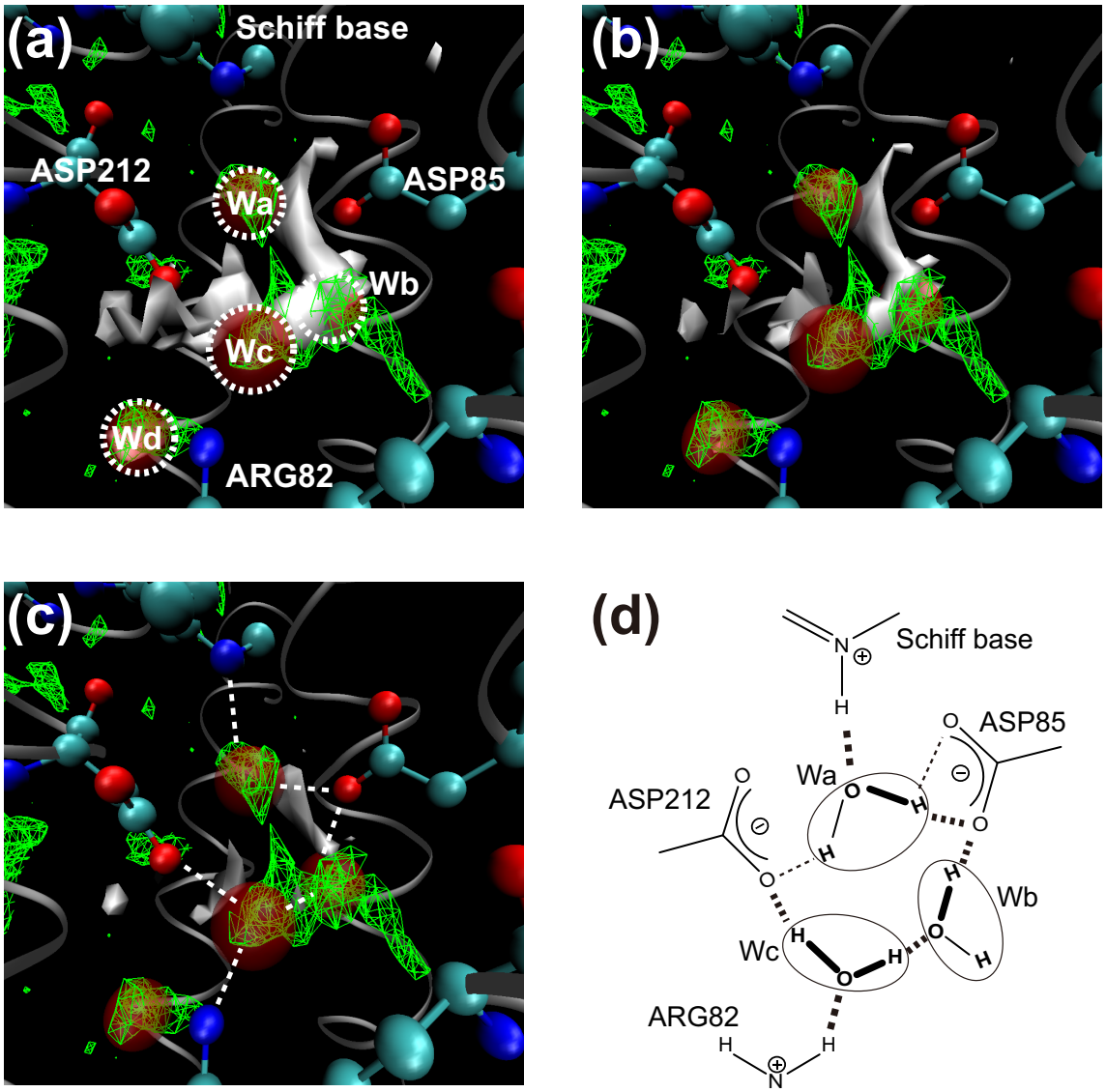


Fig. 2. Yokogawa *et al.*

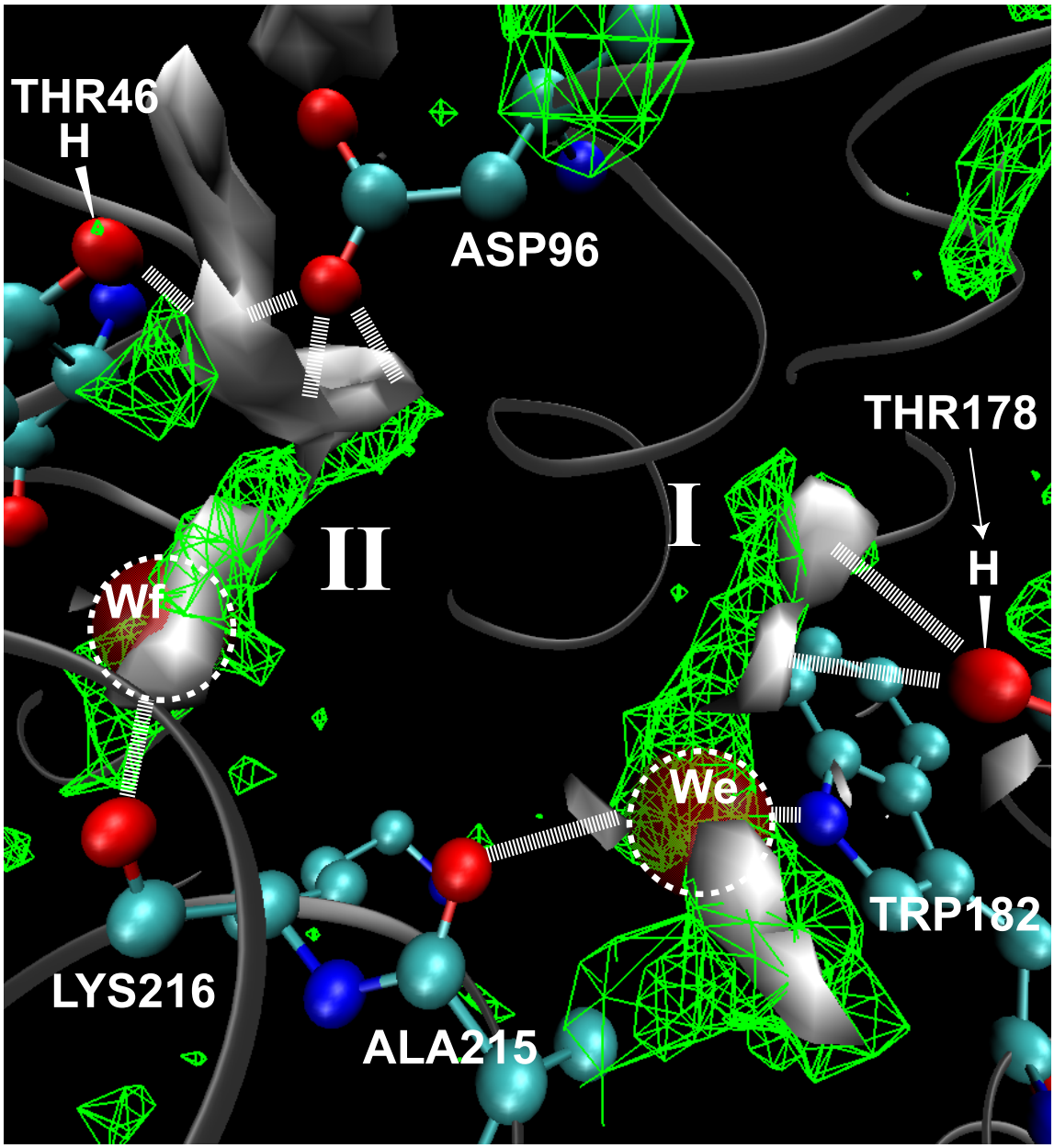


Fig. 3a. Yokogawa *et al.*

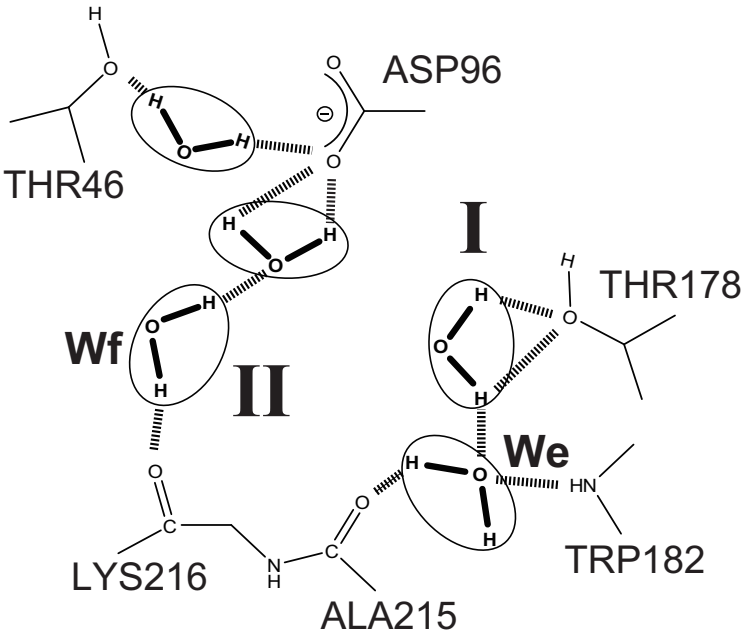


Fig. 3b. Yokogawa *et al.*